The causal analysis of operant selection by means of counterfactual models

Matthias Borgstede, Patrick Anselme***

*Corresponding author:

University of Bamberg

Markusplatz 3

D-96047 Bamberg, Germany

Tel.: +49 (0)951 863 2281

matthias.borgstede@uni-bamberg.de

**Ruhr-University Bochum

D-44780 Bochum, Germany

patrick.anselme@rub.de

The causal analysis of operant selection by means of counterfactual models

Abstract

We present a new methodology to partition different sources of behavior change within a selectionist framework based on the Price equation – the Multilevel Model of Behavioral Selection (MLBS). The MLBS provides a theoretical and formal background to disentangle behavior changes due to selection effects (e.g., “reinforcement”) from behavior changes due to non-selection effects (e.g., “skill acquisition”).

The MLBS allows for the construction of theory-based causal models that can be used to disentangle different sources of behavior change, even if they cannot be isolated experimentally. Using counterfactual analysis, we show how to quantify the causal effects of selection and non-selection sources of behavior change for a given subject. We further demonstrate how these effects may be tested for statistical significance using a simulation-based permutation test framework. The proposed method is especially useful in situations where confounding variables are difficult to control, such as changes in satiation, motivation or skill within individuals, or environmental fluctuations as they frequently occur in naturalistic settings.

Keywords: *operant selection, Price equation, covariance based law of effect, multilevel model of behavioral selection, causal inference, counterfactual model*

1 Introduction

The quantitative analysis of behavior traditionally focuses on establishing functional relations between behavior and the contextual factors in which the observed behavior occurs, such that one behavioral measure or more (e.g., lever presses or key pecks) is expressed as a mathematical

function of one independent variable or more (e.g., delay or amount of food delivery). For example, in matching analysis, the ratio between two observed behavioral measures, such as the number of pecks at two different keys, is described as a function of the ratio between two experimental feedback measures, such as the number of food pellets received contingent on pecking at each key (Herrnstein, 1961). Functional relations help structure the observed data, predict future behavior, and advance new ways to account for behavior. Furthermore, all other factors being held constant in an experiment, a functional relation between an independent and a dependent variable allows us to infer a *causal effect* of the independent variable (Menzies & Price, 1993).

However, causal inference is delicate when behavior is analyzed as a function of *contingent* consequences, such as food delivery according to a specified schedule of reinforcement. Although contingent food delivery may be a cause of feeding related behavior such as pecking, any contingency logically implies that the selected behavior is also a cause of food delivery. Therefore, the relation between behavior and food delivery constitutes a closed causal loop (Baum, 2018), making it difficult to isolate the causal effects of food delivery. Moreover, the presence of food usually raises the likelihood of feeding related behavior even if the food is not contingent on behavior (Baum, 2015). Contingencies, by design, introduce a causal path from the behavior to its consequences. Therefore, functional relations between contingent consequences and behavior are always limited with regard to causal interpretations of the model parameters. For example, as pointed out by Staddon (2020), the matching law is actually a functional relation between two *dependent* variables – a ratio of behaviors and a ratio of contingent consequences. Therefore, one cannot interpret the parameters of the fitted matching equation in terms of causal effects.

Causal inference appears even more difficult when behavior is studied under non-equilibrium conditions. For example, behavioral dynamics may change over time due to gradual depletion in food density, satiation, fatigue or skill acquisition. Such confounding variables are difficult to account for because they mainly depend on the animal's behavior and thus cannot be directly controlled by

the experimenter. Although some of these confounds may be indirectly controlled by special experimental protocols (e.g., percentile schedules), it is not always possible to implement these designs. For example, a percentile schedule requires to adjust the experimental feedback dynamically to the animal's behavior, which may interfere with the experimenter's aims, when the main interest is to investigate stable state behavior under constant conditions.

Finally, experimental control conditions are not always available due to incomplete or sub-optimal data, as it is often the case when data from previous experiments are re-analyzed or when data is collected in natural environments. For example, imagine a pigeon searching for food in an unfamiliar, open-field apparatus. We may quantify its foraging behavior by measuring the number of observed pecks over several successive time intervals. Say that one food patch contains significantly more food than the other patches. The pigeon's pecking behavior will most probably change over the course of time as a function of its environmental consequences, such as differential capture rates. At the same time, the environmental feedback will change as a function of pecking behavior because the food sources will be gradually depleted. However, it is difficult to predict the direction of change, because behavior may be influenced by several factors simultaneously. Because feeding is contingent on pecking, we might expect pecking behavior to eventually be selected towards higher peck rates at the richer patch (traditionally, such an effect would be called *reinforcement learning* or *operant conditioning*). However, an increased peck rate at the richer patch may also result from the more frequent food encounters at the richer patch that can be expected to induce further pecking nearby. Pecking behavior may also be affected by other factors that are completely unrelated to selection, such as visual detectability of food (Anselme et al., 2018; Anselme et al., 2022). Furthermore, the pigeon may eventually acquire skills that facilitate food retrieval, which in turn might reduce overall pecking. Factors like satiation or fatigue may also reduce pecking over time. Depending on the specific composition of situational and motivational factors, pecking at the rich food patch may either increase due to selection (e.g., reinforcement or induction), decrease due to

non-selection factors (e.g., satiation, fatigue or skill acquisition) or remain approximately constant if both are equally effective.

The above example illustrates that it can be difficult to infer and quantify the causal structure of adaptive behavior based on the observed behavior change alone. In some cases, the effects of confounding variables may be falsely interpreted as indicators of contingency-based selection of behavior. In other cases, selection may in fact be effective but counterbalanced by other (non-selection) sources of behavior change. Although experimental control of confounding variables is certainly the gold-standard in behavior analysis, sufficient control may sometimes be hard to achieve. Especially in field studies and applied settings, the causal loop that is intrinsic to adaptive behavior, the difficulty of appropriate experimental controls (especially in more complex and natural environments), and the dynamic nature of behavior pose severe challenges for the analysis and interpretation of behavioral data.

In this article, we propose that counterfactual reasoning by means of theory-based causal models may contribute to the existing toolkit of behavior analysis by simulating adequate control conditions when they cannot be realized experimentally. In particular, we aim to apply the counterfactual approach to causal inference in the context of operant behavior. We provide a coherent theoretical background for operant behavior that builds on the concept of selection by consequences (Skinner, 1981). We start with an introduction to the selectionist account of operant behavior and its formalization within the Multilevel Model of Behavioral Selection (MLBS) (Borgstede & Eggert, 2021). Second, we use the MLBS to develop a methodological approach that allows for the quantification of the causal effects of operant selection by means of counterfactual reasoning (Lewis, 1973). The method also allows for the calculation of p-values to test the selection effect for statistical significance on the level of single subjects. Third, we apply our methodology to quantify the causal effects of operant selection in the presence of other (non-selection) sources of behavior change using training data from a published pigeon experiment. We thereby show that operant

selection may be causally effective even when counterbalanced by non-selection factors, and discuss the implications with regard to behavioral selection theory and potential practical applications of the method.

2 Behavioral Selection Theory

Reinforcement has repeatedly been characterized as a process that works analogous to natural selection (Broadbent, 1961; D. T. Campbell, 1956; Gilbert, 1970; Herrnstein, 1964; Palmer & Donahoe, 1992; Pringle, 1951; Skinner, 1966; Thorndike, 1900). Whereas in natural selection, species adapt to the environment as a result of the fitness consequences of inheritable traits, reinforcement consists of individuals adapting to specific contexts as a result of the reinforcing consequences of repeatable actions. Skinner (1981) proposed that both processes should be subsumed under the common explanatory mode of *selection by consequences*.

The idea that reinforcement may be a selection process was later formalized by Baum (2017) and further developed by Borgstede and Eggert (2021), who integrated individual-level behavioral selection with population-level natural selection within the MLBS. The MLBS builds on the abstract description of selection processes by means of the Price equation (Price, 1970, 1972). The Price equation describes selection by means of the covariance between the individual values of a quantitative character (e.g., an inheritable trait, such as size) and individual evolutionary fitness (i.e., the contribution of an individual to the future population). A positive covariance is associated with a positive change in mean character value (i.e., selection results in a higher average character value), whereas a negative covariance indicates a negative change (i.e., selection results in a lower average character value). In other words, those character values that are statistically related to higher evolutionary fitness, are selected and thus alter the population average. The Price equation provides a mathematical description of this process that is independent of the specific mechanisms of variation, selection, and transmission (Luque, 2017).

In the Price equation, all other sources of change (those that do not refer to selection) are subsumed in a residual term. This term has different interpretations depending on the context. For example, when applied to the evolutionary change of gene frequencies, the residual term captures the effects of imperfect transmission (i.e., mutation and recombination). When applied to phenotypic change over generations (e.g., body weight or behavioral traits), the term may also capture environmental factors influencing the phenotype (Luque & Baravalle, 2021).

In the MLBS, the Price equation framework is applied to a population of individuals that vary in a certain quantitative behavior (such as average number of pecks emitted in the presence of certain environmental cues). On the population level, the covariance term refers to the effects of natural selection on average behavior tendencies of the population, and the residual term captures the average change within individuals. The MLBS is mainly concerned with this latter within-individuals change. Following the rationale that individual changes in behavior can be explained through selection by consequences, the MLBS extends the Price equation by applying the same covariance principle on the within-individuals level. Here, the population average in behavior is not calculated over different individuals, but over recurring instances of the same context (e.g., experimental trials). Behavior change, such as an increase or decrease in key pecking rate, averaged over a longer period, is described according to the covariance partitioning from the Price equation. However, the criterion of selection is not individual fitness itself (in terms of a direct contribution to the future population), but statistical predictors of individual fitness. In other words, on the within-individual level, behavior is not selected by means of reproduction or survival (in fact, if the individual dies, all of its behavior immediately ceases), but by events that predict expected change in evolutionary fitness (Borgstede, 2020). For example, food is generally a positive predictor of evolutionary fitness because it raises the probability of survival and, thus, future reproduction. Conversely, physical threat is a negative predictor of evolutionary fitness because it lowers the chances of survival and future reproduction. The concept of a fitness predictor is largely equivalent to what Baum (2012) calls a “phylogenetically important event (PIE)”. The main difference is that, in the MLBS, fitness

predictors are defined on purely statistical grounds, whereas Baum emphasizes the causal effects of PIEs on evolutionary fitness.

The conceptual framework of the MLBS allows us to describe behavioral selection on the individual level by means of a within-individuals covariance between behavior in recurring contexts and its consequences in terms of statistical fitness predictors.[1] Similar to the process of natural selection, those behaviors that covary with events that signal a gain in expected evolutionary fitness are selected, which in turn changes the average behavior of the individual. Following the MLBS, the amount of behavior change due to selection corresponds to the covariance between behavior and fitness predictor, weighted by the slope of the fitness function of the fitness predictor. Formally, the amount of behavior change due to within-individual selection can be expressed by the following equation (Borgstede & Eggert, 2021):

$$\Delta_S b = \beta_{wp}\mathrm{Cov}(b, p) \tag{1.}$$

Here, b designates a quantitative behavior (such as the number of observed key pecks in a recurring context) and $\Delta_S b$ corresponds to the within-individual change in average behavior due to selection. The value p is a quantitative fitness predictor and $\mathrm{Cov}(b, p)$ is the covariance between the behavior and the fitness predictor. Finally, β_{wp} is the slope of the function relating the fitness predictor p to the actual fitness w.

Just like the original Price equation, the MLBS introduces a non-selection term that captures all other sources of behavior change. Designating the overall change in average behavior, Δb is the within-individuals behavior change which can now be expressed as the sum of a selection term and a non-selection term δ:

[1] Note that statistical fitness predictors are not limited to events that directly increase fitness (such as mating or feeding), but also includes indirect predictors of fitness, such as information about consistent cues for the availability of food, e.g., Anselme (2022); Fortes et al. (2016); McDevitt et al. (2016); Zentall (2016). Parameters that control information seeking might be relevant predictors of evolutionary fitness, even if their influence may be less direct (e.g., Inglis et al. (1997); cf. also Borgstede (2021)).

$$\Delta b = \Delta_S b + \delta \tag{2.}$$

This equation is called the covariance based law of effect (CLOE).[2] The CLOE can be regarded as a fundamental principle of behavior analysis in that it captures the essence of behavioral selection by partitioning the overall change in behavior into a selection component and a non-selection component (Borgstede & Luque, 2021). The selection component largely corresponds to what is traditionally called reinforcement (i.e., contingency-based effects of behavioral consequences), whereas the non-selection term subsumes all other sources of behavior change (e.g., motivational factors, fatigue or skill acquisition). In the following section, we use the MLBS as a starting point for a theory-based approach to measuring selection as a source of within-individual behavior change (cf. Borgstede & Eggert, 2023a, 2023b).

3 Causal Analysis and Counterfactual Reasoning

On a theoretical level, the MLBS accounts for a great range of behavioral phenomena that are difficult to explain using traditional stimulus-response theories (cf. Borgstede, 2021; Borgstede & Eggert, 2021; Borgstede & Luque, 2021), such as the blocking effect (Kamin, 1969), operant selection by means of rare behavior (Premack & Premack, 1963) the establishment of effective contingencies through response deprivation (Timberlake & Allison, 1974), the relation between operant selection and information seeking (Berlyne, 1957; Hendry, 1965), and various deviations from matching (Davison & McCarthy, 2016).

However, quantitative estimates of behavioral selection from empirical data can be difficult to obtain (cf. Strand et al., 2022). As mentioned before, selection effects may be counterbalanced by other sources of behavior change, which is why a direct correspondence between observed behavior change and selection is not guaranteed (cf. Borgstede, 2023). Moreover, observed behavior may differ between trials due to random fluctuations of behavior, rather than selection. Therefore,

[2] See Appendix for a mathematical derivation of the CLOE from the elementary Price equation.

estimates of behavioral selection from empirical covariances between behavior and reinforcement will often only be reliable for large amounts of data. However, covariance estimates from a long time series may be biased if the process is not stationary[3] (Hamilton, 1994). Since selection introduces non-random behavior change which shifts the mean value of the behavior distribution over time, direct estimates of covariance between behavior and fitness predictors may sometimes be unreliable.

The above challenges are not unique to the theory of behavioral selection, but also occur in biological applications of the Price equation. Although the partitioning of change into selection and non-selection components is always possible on a theoretical level, empirical applications require specific models about the causes of change (Okasha & Otsuka, 2020). Okasha and Otsuka (2020) propose to use such models for causal inference using *counterfactual reasoning* (Lewis, 1973, 2000). Given a specific model, counterfactual analysis yields empirical estimates of the causal effects of the selection and non-selection components in the Price equation. In the domain of behavioral selection, a causal interpretation of the selection term in the CLOE would provide a quantitative estimate of the causal effect of operant selection.

The general idea behind counterfactual reasoning is that we attribute a causal effect to an event if we have reason to believe that, all other things being equal, the effect would not have occurred in the absence of the event (Pearl, 2013). Given this interpretation of causality, conducting a controlled experiment is an attempt to produce two exchangeable scenarios that are identical in every relevant respect, except for the event that we assume to be the cause of what we expect to observe. In between-subject experiments, we may ensure exchangeability by randomly assigning the individuals to the experimental conditions. In within-subject experiments, we may ensure exchangeability by randomly choosing the temporal order of the experimental condition for each subject (Edgington &

[3] A stochastic process is called stationary if the parameters describing the probability distribution (such as the mean and the variance) do not change over time.

Onghena, 2007). Applying counterfactual reasoning, a properly designed experiment shows us, what would have happened to an individual's behavior in condition A, if it had been assigned to condition B, instead (and vice versa).

As mentioned in section 1, it is sometimes difficult, if not impossible, to create experimental conditions that are equal in every relevant respect apart from the independent variable. Nevertheless, we may still apply the method of counterfactual reasoning to justify causal inference, as long as we make our inference conditional on a specific model. In other words, given a theoretical model, we may use empirical observations to estimate *what would have happened if certain parameters had been constant*. In this sense we may justify a claim like "the gravitational field of the sun causes the earth to move on its observed trajectory", although it is technically and physically impossible to change the gravitational field of the sun in a controlled experiment to verify our causal claim. We may, however, justify this causal claim based on a theoretical model that describes the trajectory of planets in different gravitational fields (e.g., classical mechanics). We can apply the model to empirical observations regarding the movement of the earth and the other planets in our solar system and use these data to estimate the gravitational pull of the sun on the earth, as well as the earth's momentum. Together, these empirical estimates yield a very accurate description of the earth's trajectory. A causal interpretation can now be justified by leaving the parameterized model unchanged, except for the parameters that correspond to the gravitational field of the sun. If we set the sun's gravitation equal to zero (in our model!), we may then calculate what would have happened to the earth without the sun's gravitational field. The difference of these two trajectories – the actually observed one and the counterfactual one – quantifies the causal effect of the sun's gravitational field.

Similar to the above physical example, we can justify theory-based causal inference using counterfactual reasoning in behavior analysis. Such a theory-based approach would be especially helpful in situations where it is difficult or impossible to justify a causal interpretation on

experimental grounds alone. Therefore, here we propose theory-based approach to infer the causal effects of operant selection from observed behavior change by means of the above described MLBS and the CLOE.

The CLOE describes change due to reinforcement by means of a covariance term, such that a higher covariance between behavior and a fitness predictor yields a higher selection and a lower covariance yields smaller selection. To apply the counterfactual method outlined above, we need to use the MLBS in the specific conditions that generated the observed data. For example, the theoretical covariance between a quantitative behavior and a quantitative fitness predictor can generally be obtained from the feedback function of a reinforcement schedule (cf. Borgstede & Luque, 2021). Similarly, the quantitative behavior itself (and, consequently, behavior change) can be observed over several repeated experimental trials. We can then use the corresponding model parameters to calculate what would have happened in the absence of selection, conditional on the specified model. In other words, assuming that individual behavior changes due to selection processes and that these processes are formally captured by the MLBS, we can calculate the counterfactual behavior by leaving everything in the model unchanged, except the part that is responsible for selection to occur (cf. Okasha & Otsuka, 2020). Practically, this means to apply a minimal change to the model parameters, such that the covariance term becomes zero (because a zero covariance implies zero selection).

Mathematically, there are several ways to ensure that the covariance term in the CLOE is zero. However, the most plausible candidate for our counterfactual model is certainly that the fitness predictor (e.g., the amount of food received) is set equal across trials, i.e., equal to the average value of the fitness predictor over all observed trials. For example, if we conduct an experiment with two subsequent trials, we may observe behavior that yields 10 food items in trial 1 and 20 food items in trial 2. The counterfactual model would use the actual feedback functions from the two trials, respectively, to calculate the behavior that would have resulted in the two trials given the individual

had continued doing what it was doing, until the average of 15 food items had been received. Conditional on the model, the difference between these two counterfactual values for the behavior in trials 1 and 2 corresponds to the behavior change that we would have observed if there had been no behavioral selection. Comparing the change with and without selection, we may calculate an empirical estimate of the causal effect of selection by taking the difference between the counterfactual change and the actually observed change. The same rationale applies if we want to calculate the causal effect of specific non-selection sources of behavior change. If, for example, we suspect a causal effect of skill acquisition, we would set the corresponding part of the non-selection term in the CLOE to zero. Since skill effects naturally translate to increases in success rate of the target behavior, we may calculate the amount of behavior change that we would have expected in the absence of skill acquisition by holding the success rate of the observed behavior constant in the model. For example, if an individual requires an average of ten attempts for a successful response in the first trial but only an average of five attempts in the second trial, we would substitute the slopes of the actual feedback functions from both trials with the slopes given a constant success rate (i.e., the geometric mean of the two observed success rates) to obtain the amount of pecks that would have yielded the observed number of food items in trials one and two.

The proposed method can be summarized by the following steps. First, describing the experimental scenario in terms of behavioral selection using the MLBS. Second, using the model to calculate the amount of behavior change that would have occurred in the absence of selection, or in the absence of a non-selection factor, respectively. Third, subtracting the behavior change obtained from these counterfactual conditions from the total behavior change observed between trials.

The counterfactual methodology provides a simple way to infer causal effects of operant selection from observed behavior change, even if confounding variables are known to be effective. In the following section, we will fully develop the method by applying it to an actual empirical data set. We will show how empirical estimates of the causal effects of selection and non-selection components

of operant behavior can be obtained at the level of individual learning, and how these individual causal effects can be evaluated, compared and tested for statistical significance using a Monte-Carlo simulation approach.

4 Exemplary Application: Behavioral Selection Between Learning Trials

We demonstrate how the counterfactual analysis of selection effects described above can be implemented in an empirical study using a minimal example for illustrative purposes. Our method is applied to the data from two training trials (first and last days) of a published behavioral experiment involving pigeons (Anselme et al., 2022). The focus of the main experiment was to investigate the effects of signals for the availability of food items in the holes of a board on foraging behavior. Here, we focus on the training trials administered prior to the main experiment, consisting of two conditions only (“no food” vs. “guaranteed food” at the beginning of a trial). The question we address is, whether the pigeons actually learned to distinguish between the food and the no-food conditions between training trials (i.e., if behavioral selection was effective). Additionally, we investigate the potential contribution of non-selection sources of behavioral change, such as increased pecking skills.

4.1 Experimental Apparatus and Data Acquisition

To test our model, we exploited some video data from a published study (Anselme et al., 2022). Here, we only provide the methodological details that are relevant to understanding our analyses. Nine naïve pigeons were maintained at 85-90% of their free-feeding body weight to motivate them to eat in the task.

The pigeons were tested in a rectangular wooden box whose floor was a horizontally removable plate of wood (120 cm length × 70 cm width × 2 cm height), perforated with holes (1.5 cm diameter and 2 cm depth). The foraging board contained 60 holes organized as 6 rows of 10 holes regularly

spaced. The board was covered with a black plastic tape with a cross cut above each hole to create an opening, which allowed the pigeons to access the food items while being unable to visually detect their presence from a distance (Figure 1). Specific stimuli (green and red; 21 cm length × 14.5 cm width) were used to signal the consistent presence or absence of food per hole in one area. The two areas were separated by means of a colored strip glued on the plastic tape, dividing the board in two equal left and right areas of 30 holes each from the entrance compartment. Each session was recorded with an external camera, placed above the apparatus.

In each of the 30 holes of one area (left or right, counterbalanced across trials within the same individuals), we positioned one food item (corn, green pea, yellow pea, or sunflower) and this area was associated with one discriminative stimulus (red or green, counterbalanced across individuals) placed on each wall (4 cm above the floor level), from the first to the last trial. For a given individual, a stimulus location was counterbalanced across trials. The 30 holes of the adjacent area remained empty and were associated with the other stimulus placed on each wall. Of note, the pigeons were initially trained for 3-4 days with uncovered holes, each containing one grain, such that they could see the food in the holes on inspection. No discriminative stimulus was used at this stage. After this initial training, the pigeons were trained as reported above for 4 consecutive days with covered holes.

Data Extraction

Data were collected on manual counting (food items consumed and number of pecks per area for each trial). Determining whether a peck at a hole was successful (food item grasped) or not was mostly impossible from the videos, so that peck is not synonymous with item consumed. A peck simply meant a vertical downshift of the pigeon’s head above a hole. We considered a pigeon to be positioned in a given area if its head was in this area—because its body could be in one area and yet pecking in the adjacent one. Sometimes, the pigeon missed a grain (picked it up and lost it), so that it

rolled on the board. In the attempt to get it, the pigeon could cross the demarcation line between two areas. A peck given outside of a hole, even to get a missed grain, was not counted.

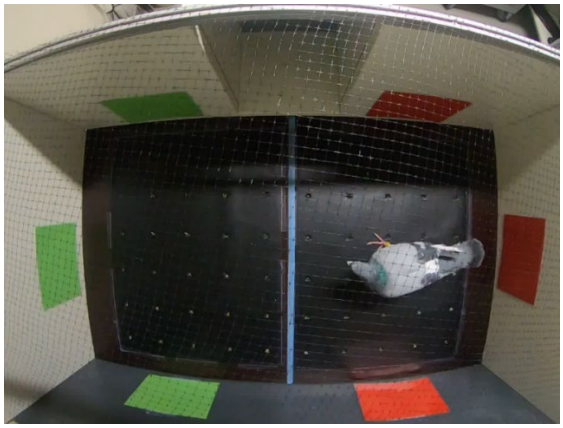

Figure 1: Experimental setting. Pigeons were put in a 120cmx70cm experimental chamber with 60 covered holes. In the food region, each hole contained a grain. In the no-food region, the holes were empty.

4.2 Statistical Analyses

We focused on the observed change in individual pecking behavior between trials and its relation to individual capture rate. The primary measures used in the analysis were the number of pecks emitted while staying at the food region during each trial and the number of food items retrieved during each trial. As a derived measure, we calculated individual success rates for each trial by dividing the number of retrieved food items by the number of pecks for each pigeon and each trial. Since pigeons are known to forage systematically, thereby avoiding sites that they already exploited (Baum, 1987), the expected number of food items per peck only depends on the average pecking success, yielding a feedback function that is linear until all grains are retrieved (for higher numbers of pecks, the slope of the feedback function is zero). Assuming that the pecking success remains approximately constant within trials, the slopes of the individual feedback functions for each trial equal the pigeon's average success rates during the respective trials (i.e., the number of pecks at the food region divided by the number of retrieved food items during a trial).[4]

[4] The analysis would be equally possible for random foraging behavior. In this case, the slope of the feedback function would also depend on the number of food items that have already been retrieved, yielding a negatively accelerated change in expected feedback. Re-analysis of the data under the assumption of random foraging changed the quantitative estimates of the MLBS but did not change the overall qualitative patterns or the group level effects.

All three variables were compared between the first trial (day 1) and the last trial (day 4) of the training phase. The respective change values were tested for significant deviations from zero ("no change") using two-sided exact permutation tests for paired data to compute p-values for increases in skill and effort, respectively (Edgington & Onghena, 2007).

To estimate the causal effects of selection and non-selection between trials, we constructed two counterfactual null conditions from the MLBS. The null conditions were obtained by parameterizing the MLBS under the assumptions that (a) operant selection due to contingent feedback was absent and (b) non-selection change due to skill acquisition was absent. The selection null condition was realized by keeping the success rates in the first and the last trials unchanged and fixing the number of retrieved food items equal to average food retrieval obtained from the arithmetic mean of the actual numbers of retrieved food items in both trials. The non-selection null condition was realized analogously by keeping the numbers of retrieved food items in the first and last trials unchanged and fixing the slope of the feedback function (i.e., the success rate) to the average slope obtained from the geometric mean of the observed success rates in both trials.

We then used the inverse of the feedback function to obtain the expected number of pecks required to obtain the respective amounts of food obtained from the counterfactual null conditions. In the selection null condition, we calculated the number of pecks it would have taken each pigeon to retrieve an average number of food items given the actual slopes of the feedback function (i.e., the observed success rates). In the non-selection null condition, we calculated the number of pecks it would have taken each pigeon to retrieve the actual numbers of retrieved food items given an average success rate (i.e., constant slope of the feedback function across trials). We then compared the actually observed changes in pecking behavior to the change that we would have expected under the two counterfactual null conditions. Panels a) and c) of Figure 2 illustrate the construction of the counterfactual conditions using the data from pigeon 97. The gray circles in both panels mark the observed numbers of pecks and retrieved food items (i.e., 212 pecks for day 1 and 149 pecks for

day 4). The gray arrows indicate the observed change in pecking behavior (-63 pecks). The black solid lines indicate the feedback functions for the first trial (day 1) and the last trial (day 4), with their slopes being equal to the ratio of retrieved food items per peck in each trial, respectively (i.e., 0.05 grains per peck for day 1 and 0.13 grains per peck for day 4). The black circles mark the expected numbers of pecks and retrieved food items under the respective counterfactual model and the black arrows indicate the expected change in pecking behavior given the respective counterfactual models. The red dashed lines illustrate the constraints applied to the counterfactual models in order to set selection (panel a) and non-selection (panel c) equal to zero. In the selection null condition (panel a), the dashed red line marks the arithmetic mean of the observed numbers of retrieved food items (i.e., 15 grains). The counterfactual amounts of pecking are simply the points where the dashed red line intersects with the observed feedback functions of the two trials (i.e., 319 pecks for day 1 and 113 pecks for day 4). The resulting counterfactual change has the same direction as the observed change but is much larger (-206 pecks). This means that we would have expected a much stronger decrease in pecking than the observed one if no selection had taken place.

In the non-selection null condition (panel c), the dashed red line marks the average feedback function (the slope is equal to the geometric mean of the observed success rates, i.e., 0.08). The corresponding counterfactual amounts of pecking for days 1 and 4 were obtained by horizontally projecting the observed values (gray circles) onto the red dashed line (i.e., 126 pecks for day 1 and 251 pecks for day 4). Thus, the resulting counterfactual change in pecking pointed in the opposite direction as the observed change and was larger than the observed change (i.e., 125 pecks). This means that we would have expected an increase (rather than the observed decrease) in pecking if the success rates of pecking had been identical across trials.

In both counterfactual conditions, the estimated causal effects of selection and non-selection were obtained by subtracting the counterfactual behavior change from the observed behavior change. For pigeon 97, the estimated selection effect was thus the difference between the observed change (i.e.,

-63 pecks) and the expected change given the counterfactual model without selection (i.e., -206 pecks), which resulted in a value of 143 pecks. The estimated non-selection effect was equal to the difference between the observed change (i.e., -63 pecks) and the expected change given the counterfactual model without skill acquisition (i.e., 125 pecks), which resulted in a value of -188 pecks.

Following a bottom-up generalization strategy (Borgstede & Scholz, 2021), we tested whether the individual selection and non-selection terms were significantly different from zero for each pigeon separately. To calculate individual p-values, the parameterized counterfactual models were simulated for 10,000 pairs of trials for each pigeon. In every iteration of the simulation, the number of pecks needed to retrieve the amount of food that would have been expected under the counterfactual model in each trial was used to calculate the corresponding counterfactual behavior change. For example, pigeon 97 retrieves 10 grains with a success rate of 0.05 in the first trial. Since the success rate is an estimate of the probability of food retrieval for a single peck, the simulation simply consists in repeatedly drawing from a random binomial variable with a success probability of .05 until 15 successes are reached. The distribution of the simulated behavioral change values from the counterfactual model resembles the sampling distribution from which the actually observed behavior change would have been drawn under the null hypothesis that the observed behavior change was generated in the absence of selection or non-selection, respectively . The quantiles of this sampling distribution were used to calculate one-sided p-values for each pigeon (compare Figure 2, panels b and d). If a p-value yielded a significant result at the 5%-level, the null hypothesis was rejected in favor of the alternative that the observed change was at least partly caused by selection or non-selection, respectively.

In addition to the individual-level p-values, we performed two-sided exact permutation tests for paired data on all observed and derived measures, thereby testing against the null hypothesis that individual change values from the first to the last trial were zero for the respective variables.

All statistical analyses and simulations were conducted using the software R, version 4.2.2 (R Core Team, 2022).

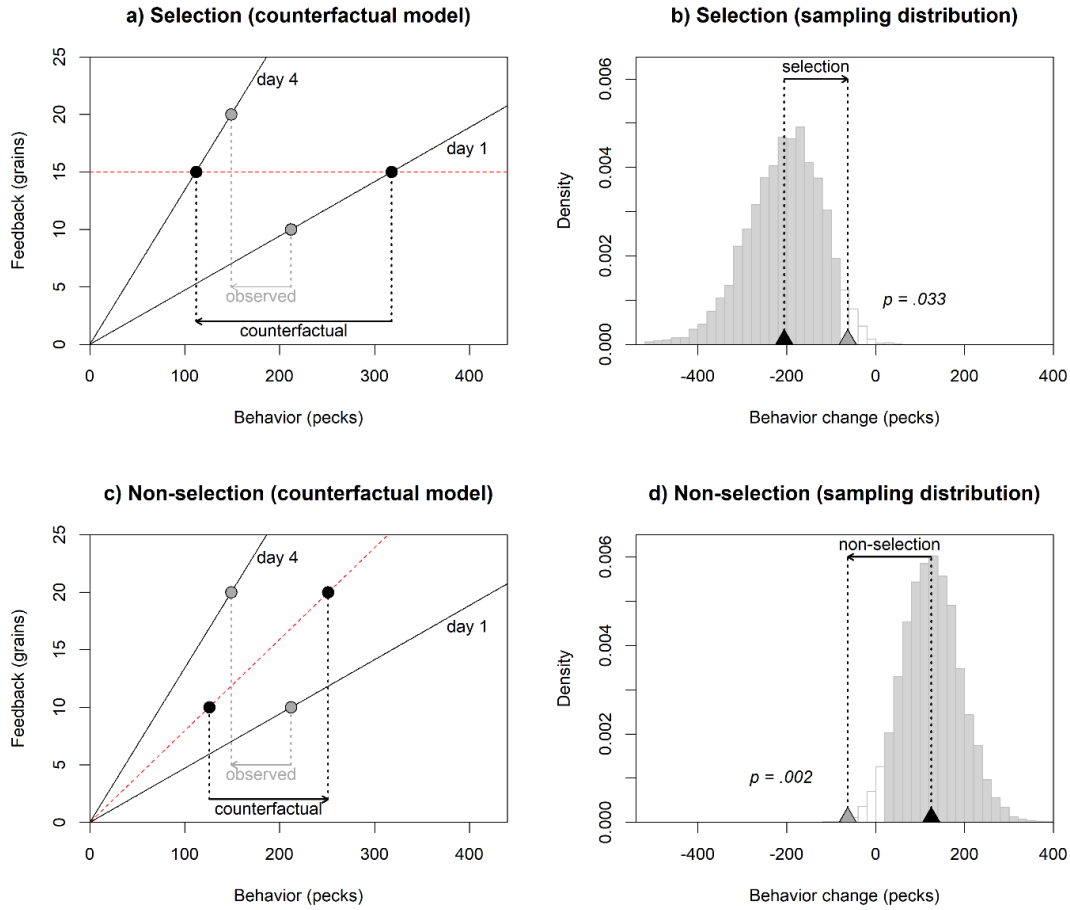

Figure 2: Counterfactual models and sampling distributions. Panels a) and c) show the feedback (food) as a function of behavior (pecking) for days 1 and 4. The red dashed lines indicate the parameter restrictions of the counterfactual models for the selection and the non-selection conditions, respectively. Panels b) and d) depict the corresponding sampling distributions under the null hypothesis that the counterfactual model was the data generating model. The lighter shades of the distributions indicate the areas of rejection. Data obtained from Pigeon 97.

4.3 Results

Table 1 summarizes the pigeons' behavior during the first and the last trial of training, respectively. The main behavioral metric was the absolute number of pecks emitted at the food region in each trial. The average number of pecks at the food region during the first trial was 230.11 (*SD* = 126.99). The average number of pecks at the food region during the last trial was 190.44 (*SD* = 94.12). Thus, average number of pecks at the food region *decreased* over training trials, although the paired samples permutation test for peck difference was not statistically significant. Despite an overall

reduction in pecking, the average number of retrieved food items *increased* from the first trial (*M* = 13.67, *SD* = 6.2) to the last trial (*M* = 20.0, *SD* = 6.15). The observed changes in pecking and food retrieval are illustrated by Figure 3a and 3b, respectively.

Dividing the number of obtained food items per trial by the number of pecks per trial for each pigeon, individual success rates were obtained. As could be expected from the observed changes in pecking and food retrieval, average success rate (captured food items per peck) increased over trials. However, the increase in success rate was not statistically significant (p = .086).

Table 1: Comparison of first and last training trial (day 1 and day 4) with respect to the number of pecks at the food region, the number of retrieved food items and the food items retrieved per peck while in the food region.

	behavior (pecks)		**feedback (grains)**		**success (grains per peck)**	
	day 1	**day 4**	**day 1**	**day 4**	**day 1**	**day 4**
P17	233	87	20	23	0.09	0.26
P73	325	340	14	17	0.04	0.05
P91	456	266	22	29	0.05	0.11
P96	237	117	13	18	0.05	0.15
P97	212	149	10	20	0.05	0.13
P103	44	335	1	20	0.02	0.06
P118	116	104	8	7	0.07	0.07
P519	88	199	18	28	0.20	0.14
P534	360	117	17	18	0.05	0.15
M	230.11	190.44	13.67	20	0.07	0.12
SD	126.99	94.12	6.2	6.15	0.05	0.06

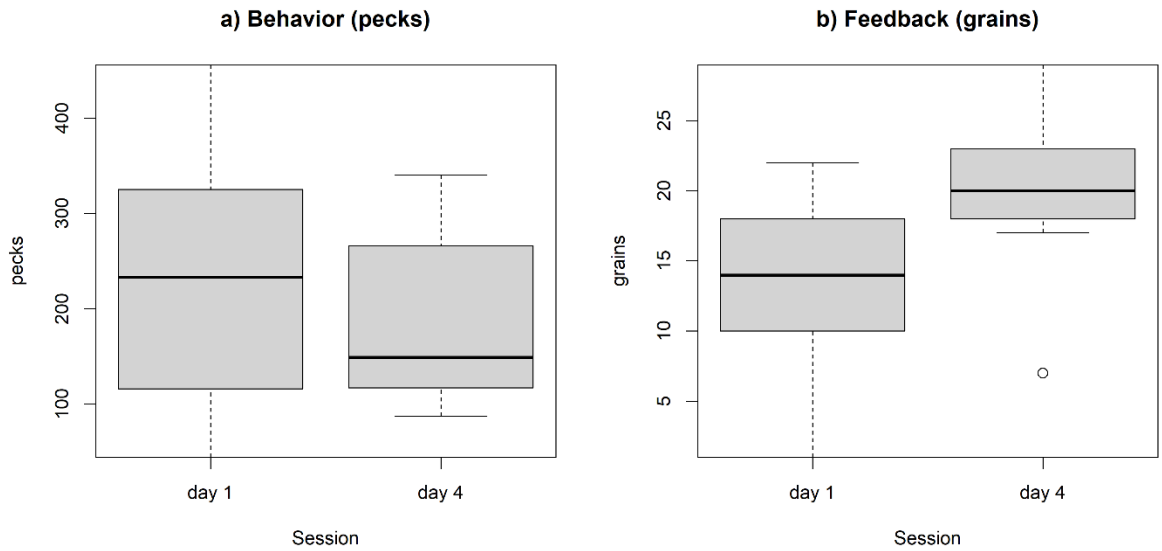

Figure 3: Changes in a) behavior (number of pecks at food region) and b) feedback (number of retrieved food items) from the first day of training to the last day of training.

Table 2 shows the results from the counterfactual analysis. In accordance with the descriptive data, the observed change in pecking at the food region, Δb, was negative for all but three individuals (pigeons 73, 103 and 519), with an average value of -39.67 pecks (*SD* = 165.09 pecks). The estimated selection effect from the counterfactual analysis was positive for all but one individual (pigeon 118). The individual p-values yielded significant selection for only two individuals (pigeons 97 and 103). Nevertheless, the average selection effect of 116 pecks (*SD* = 174.14 pecks) was significant on the 5% level ($p = .023$). The estimated non-selection effect from the counterfactual analysis was negative for all but two individuals (pigeons 118 and 519). The individual p-values indicated a significant amount of non-selection (i.e., skill acquisition) for six individuals. The average non-selection effect from the counterfactual analysis was -143.33 pecks (*SD* = 118.01 pecks), which was also significant on the 5% level ($p = .012$). Subtraction of the sum of the selection and non-selection estimates from the observed changes yielded the residuals (i.e., the unexplained partition of behavior change) with an average of -12.33 pecks (*SD* = 14.54 pecks). The unexplained behavior change was significantly different from zero ($p = .012$), indicating that there might be other sources of behavior change over and above contingent food delivery and skill acquisition. The results from

the counterfactual analysis are illustrated in Figure 4, which shows that behavioral selection was generally positive, despite a negative overall change in behavior.

Table 2: Partitioning of total behavioral change (pecking at food region) from first to last training day. The observed behavioral change (Δb) is expressed as the sum of the estimated selection effect ($\Delta_S b$) and the estimated non-selection effect (δ). The residual change equals $\Delta b - (\Delta_S b + \delta)$, respectively. Individual p-values were calculated using a Monte-Carlo approximation to the sampling distribution under the null hypothesis that the true selection effect is zero ($H_0: \Delta_S b = 0$) and under the null hypothesis that the true non-selection effect is zero ($H_0: \delta = 0$). The corresponding group-level p-values were obtained from two-sided exact permutation tests.

	observed change (pecking)	selection (reinforcement)	non-selection (skill acquisition)	residual (unexplained)
P17	-146	27	-166*	-7
P73	15	66	-49	-2
P91	-190	111	-287*	-14
P96	-120	68	-174*	-14
P97	-63	143*	-188*	-18
P103	291	563*	-225*	-47
P118	-12	-15	2	1
P519	111	60	52	-1
P534	-243	21	-255*	-9
M	-39.67	116*	-143.33*	-12.33*
SD	165.09	174.14	118.01	14.54

Note: Individual p-values from Monte-Carlo simulations, group level p-values from permutation tests
* statistically significant on 5%-level

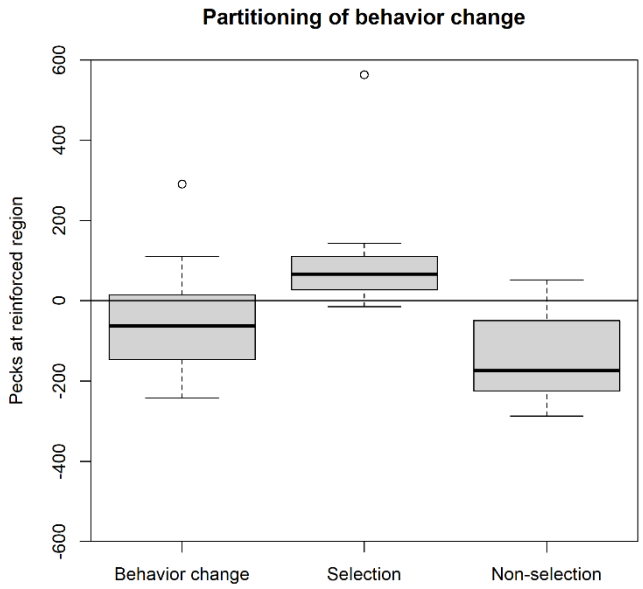

Figure 4: Partitioning of behavior change (pecks at reinforced region) into selection (reinforcement) and non-selection (skill acquisition) components using the Covariance Based Law of Effect (CLOE).

5 Conclusion

In this paper, we proposed a new method to quantify the causal effects of operant selection by means of counterfactual reasoning. The method builds on a formal theory of operant selection, the Multilevel Model of Behavioral Selection (MLBS), which allows for partitioning the observed changes in average behavior into a selection component and a non-selection component. Applying the MLBS to empirical data, we showed how the causal effects of selection and non-selection can be inferred, even if confounding variables are effective. The rationale was to use the MLBS to simulate what would have happened in the absence of selection or non-selection effects, respectively. The difference between the expected (counterfactual) behavior change and the observed (actual) behavior change yields estimates of the selection and non-selection components of operant behavior. We further proposed a Monte-Carlo simulation framework to statistically test against the null hypothesis that selection (or non-selection, respectively) was absent at the level of single individuals.

The proposed method was capable to separate selection effects (i.e., behavior change due to contingent food delivery) from non-selection effects (i.e., behavior change due to increased skills) in a foraging experiment. Although the observed change between training trials was negative (i.e., individuals tended to decrease pecking behavior at the food region), there was in fact strong evidence for positive behavioral selection. However, these selection effects were more than counterbalanced by strong increases in skill. In other words, the pigeons did not only learn where the food was, but also became more efficient in retrieving the food from the apparatus. These results indicate that selection and non-selection both contributed to the observed change in behavior.

We present the first empirical application of the MLBS in the context of a behavioral experiment. The experiment itself was chosen such that it is as simple as possible, to serve as a minimal example for the method proposed in this article. Of course, there are various limitations with regard to the

data because they were originally collected for a different purpose. For example, the foraging board was constructed in a way that one could not unequivocally identify the retrieval of food on the videos. Making the floor below the cover transparent might have solved this problem. However, this would have possibly enabled the pigeons to see where the food is (because of the light emerging from the holes), a situation likely to affect their foraging behavior. Despite the limitations of the experimental application, our results show that the MLBS in combination with the counterfactual modeling approach provides a feasible theoretical foundation for the experimental analysis of behavior and valuable starting point to justify theoretically meaningful quantifications (i.e., measurement) in the domain of psychology (cf. Borgstede & Eggert, 2023a, 2023b).

The proposed selection framework acknowledges that contingent food delivery cannot always be isolated as an independent variable. Moreover, the resulting methodology allows us to disentangle selection and non-selection effects, even if behavior is in a non-equilibrium state. Shifting the focus away from steady-state analysis (such as the application of the generalized matching law), the proposed method specifically applies to behavior that is changing, such as the gradual behavior change during operant training. In general, data from training trials can be messy due to various uncontrollable factors affecting behavior simultaneously, including confounding variables that change within individuals and can thus not be directly controlled by the experimenter. We think that the counterfactual analysis of behavioral selection provides the necessary tools to deal with this kind of data. We would like to stress, however, that we do *not* propose to replace existing experimental control protocols where they are available. If confounding variables can be effectively controlled in an experiment, they certainly should be. Nevertheless, the ideal of perfect experimental control is hard – and often impossible – to achieve. Therefore, we think that it is worthwhile exploring model-based approaches to at least statistically account for uncontrolled confounds where they are unavoidable.

The general methodology of model-based counterfactual analysis can be applied to many other experimental settings that involve behavior change over time – probably, there are already thousands of unused training data sets only awaiting to be analyzed. The partitioning of behavior change into selection and non-selection components may equally be applied to foraging activity or other behaviors in natural settings. Moreover, the method provides a promising framework to analyze the complexities of human behavior, such as the relation between incentives and competence gain in students' learning behavior or the interaction between environmental contingencies and verbal instructions in the formation of rule-governed behavior. We hope that this article contributes to the foundations of behavioral selection as a general theory of behavior, and encourages other researchers to put the behavioral selection perspective into practice.

6 Author Contributions

MB developed the modeling framework and the statistical methodology, analyzed the data and wrote the original draft. PA planned and conducted the experiments, collected the data and contributed to the theoretical background and the interpretation of the results. Both authors contributed to the final version of the manuscript.

7 Competing Interests

The authors declare no competing interests.

8 Data and Code Availability

The original data and R-code used in this study will be made publicly available on publication.

9 Funding

This research was supported by the Deutsche Forschungsgemeinschaft through An1067/3-1 (PA).

10 Appendix

10.1 Derivation of the Covariance Based Law of Effect

The Covariance Based Law of Effect (CLOE) is basically a recursive expansion of the elementary Price equation (Price, 1970):

$$\overline{w}\Delta\overline{b} = \mathrm{Cov}_i(w_i, b_i) + \mathrm{E}_i(w_i\Delta b_i)$$

In the Price equation, w_i designates an individual's evolutionary fitness and b_i the value of an arbitrary evolving character. $\overline{w}$ and $\overline{b}$ designate the population averages of w_i and b_i, respectively. Cov_i and E_i are the population covariance and the expected value of the population. Δb_i is the individual-level change in character value (usually thought of as change between parent and offspring) and $\Delta\overline{b}$ is the population-level change in character value.

The term $\mathrm{Cov}_i(w_i, b_i)$ captures the effects of natural selection, while the term $\mathrm{E}_i(w_i\Delta b_i)$ refers to changes in the population average of b that are not natural selection. If the time frame is chosen sufficiently small and individuals are treated as their own offspring, $\mathrm{E}_i(w_i\Delta b_i)$ captures changes that occur within individuals, such as skill acquisition or reinforcement.

In the MLBS, the fitness-weighted within-individual change, $w_i\Delta b_i$, is itself partitioned into a covariance term and an expectation term:

$$w_i\Delta b_i = \mathrm{Cov}_j\left(w_{ij}, b_{ij}\right) + \mathrm{E}_j\left(w_{ij}\Delta b_{ij}\right)$$

Whereas in the original Price equation, the covariance and expectation are taken over the individuals i of a population, in the MLBS, they are taken over a collection of recurring contexts j (so-called behavioral episodes) that are themselves nested within individuals. Consequently, the covariance term $\mathrm{Cov}_j\left(w_{ij}, b_{ij}\right)$ refers to the part of within-individual change that can be attributed to selection on the individual level (i.e., reinforcement), whereas the expectation term $\mathrm{E}_j\left(w_{ij}\Delta b_{ij}\right)$ captures all sources of within-individual change that are not selection.

Given an arbitrary fitness predictor p (e.g., food), evolutionary fitness can be predicted by a linear regression of the form $w = \beta_{wp}p + \varepsilon$. We can now re-arrange to obtain:

$$w_i \Delta b_i = \beta_{wp} \mathrm{Cov}_j\left(p_{ij}, b_{ij}\right) + \mathrm{E}_j\left(w_{ij} \Delta b_{ij}\right)$$

Defining $\Delta_S b = \frac{\beta_{wp}}{w_i} \mathrm{Cov}_j\left(p_{ij}, b_{ij}\right)$ and $\delta = \frac{\mathrm{E}_j\left(w_{ij} \Delta b_{ij}\right)}{w_i}$, we get the Covariance Based Law of Effect (CLOE):

$$\Delta b = \Delta_S b + \delta$$

11 References

Anselme, P. (2022). The optimality of “suboptimal” choice: A psycho-evolutionary perspective. In M. Krause, K. L. Hollis, & M. R. Papini (Eds.), *Evolution of Learning and Memory Mechanisms* (pp. 193–209). Cambridge University Press.

Anselme, P., Dreher, T., & Güntürkün, O. (2018). Pigeons consistently prefer easy over harder access to food: No reversal after direct dopaminergic stimulation. *Behavioral Neuroscience*, *132*(4), 293–301. https://doi.org/10.1037/bne0000249

Anselme, P., Wittek, N., Oeksuez, F., & Güntürkün, O. (2022). Overmatching under food uncertainty in foraging pigeons. *Behavioural Processes*(201), 104728. https://doi.org/10.1016/j.beproc.2022.104728

Baum, W. M. (1987). Random and systematic foraging, experimental studies of depletion, and schedules of reinforcement. In A. C. Kamil, J. R. Krebs, & H. R. Pulliam (Eds.), *Foraging Behavior* (pp. 587–605). Plenum Press.

Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior*, *97*(1), 101–124. https://doi.org/10.1901/jeab.2012.97-101

Baum, W. M. (2015). The role of induction in operant schedule performance. *Behavioural Processes*, *114*, 26–33. https://doi.org/10.1016/j.beproc.2015.01.006

Baum, W. M. (2017). Selection by consequences, behavioral evolution, and the price equation. *Journal of the Experimental Analysis of Behavior*, *107*(3), 321–342. https://doi.org/10.1002/jeab.256

Baum, W. M. (2018). Three laws of behavior: Allocation, induction, and covariance. *Behavior Analysis: Research and Practice*, *18*(3), 239–251. https://doi.org/10.1037/bar0000104

Berlyne, D. E. (1957). Uncertainty and conflict: A point of contact between information-theory and behavior-theory concepts. *Psychological Review*, *64, Part 1*(6), 329–339. https://doi.org/10.1037/h0041135

Borgstede, M. (2020). An evolutionary model of reinforcer value. *Behavioural Processes*, 104109. https://doi.org/10.1016/j.beproc.2020.104109

Borgstede, M. (2021). Why do individuals seek information? A selectionist perspective. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.684544

Borgstede, M. (2023). Fisher's fundamental theorem. In T. K. Shackelford (Ed.), *Encyclopedia of Sexual Psychology and Behavior.* Springer Cham.

Borgstede, M., & Eggert, F. (2021). The formal foundation of an evolutionary theory of reinforcement. *Behavioural Processes*, *186*, 104370. https://doi.org/10.1016/j.beproc.2021.104370

Borgstede, M., & Eggert, F. (2023a). Meaningful measurement requires substantive formal theory. *Theory & Psychology*, *33*(1), 153–159. https://doi.org/10.1177/09593543221139811

Borgstede, M., & Eggert, F. (2023b). Squaring the circle: From latent variables to theory-based measurement. *Theory & Psychology*, *33*(1), 118–137. https://doi.org/10.1177/09593543221127985

Borgstede, M., & Luque, V. J. (2021). The covariance based law of effect: A fundamental principle of behavior. *Behavior and Philosophy*, *49*, 63–81. https://behavior.org/wp-content/uploads/2022/01/BP-v49-Borgstede2_corrected.pdf

Borgstede, M., & Scholz, M. (2021). Quantitative and Qualitative Approaches to Generalization and Replication-A Representationalist View. *Frontiers in Psychology*, *12*, 605191. https://doi.org/10.3389/fpsyg.2021.605191

Broadbent, D. E. (1961). *Behaviour*. Methuen.

Campbell, D. T. (1956). Adaptive behavior from random response. *Behavioral Science*, *1*(2), 105–110. https://doi.org/10.1002/bs.3830010204

Davison, M., & McCarthy, D. (2016). *The Matching Law*. Routledge. https://doi.org/10.4324/9781315638911

Edgington, E., & Onghena, P. (2007). *Randomization Tests, Fourth Edition* (4th ed.). *Statistics*. CRC Press.

Fortes, I., Vasconcelos, M., & Machado, A. (2016). Testing the boundaries of "paradoxical" predictions: Pigeons do disregard bad news. *Journal of Experimental Psychology. Animal Learning and Cognition*, *42*(4), 336–346. https://doi.org/10.1037/xan0000114

Gilbert, R. M. (1970). Psychology and biology. *Canadian Psychologist/Psychologie Canadienne*, *11*(3), 221–238. https://doi.org/10.1037/h0082574

Hamilton, J. D. (1994). *Time series analysis*. Princeton Univ. Press.

Hendry, D. P. (1965). *Reinforcing Value of Information: NASA Technical Report No. 65-1.* University of Maryland.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267–272. https://doi.org/10.1901/jeab.1961.4-267

Herrnstein, R. J. (1964). Will. *Proceedings of the American Philosophical Society,*, *108*(6), 455–458.

Inglis, Forkman, & Lazarus (1997). Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behaviour*, *53*(6), 1171–1191. https://doi.org/10.1006/anbe.1996.0320

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296).

Lewis, D. (1973). Causation. *The Journal of Philosophy*, *70*(17), 556. https://doi.org/10.2307/2025310

Lewis, D. (2000). Causation as Influence. *The Journal of Philosophy*, *97*(4), 182. https://doi.org/10.2307/2678389

Luque, V. J. (2017). One equation to rule them all: a philosophical analysis of the Price equation. *Biology & Philosophy*, *32*(1), 97–125. https://doi.org/10.1007/s10539-016-9538-y

Luque, V. J., & Baravalle, L. (2021). The mirror of physics: on how the Price equation can unify evolutionary biology. *Synthese*, *199*, 12439–12462. https://doi.org/10.1007/s11229-021-03339-6

McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, *105*(1), 23–40. https://doi.org/10.1002/jeab.192

Menzies, P., & Price, H. (1993). Causation as a Secondary Quality. *The British Journal for the Philosophy of Science*, *44*(2), 187–203. https://doi.org/10.1093/bjps/44.2.187

Okasha, S., & Otsuka, J. (2020). The Price equation and the causal analysis of evolutionary change. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1797), 20190365. https://doi.org/10.1098/rstb.2019.0365

Palmer, D. C., & Donahoe, J. W. (1992). Essentialism and selectionism in cognitive science and behavior analysis. *The American Psychologist*, *47*(11), 1344–1358. https://doi.org/10.1037/0003-066X.47.11.1344

Pearl, J. (2013). *Causality: Models, reasoning, and inference* (Second edition, reprinted with corrections). Cambridge Univ. Press.

Premack, D., & Premack, A. J. (1963). Increased eating in rats deprived of running. *Journal of the Experimental Analysis of Behavior*, *6*, 209–212. https://doi.org/10.1901/jeab.1963.6-209

Price, G. R. (1970). Selection and Covariance. *Nature*, *227*(5257), 520–521. https://doi.org/10.1038/227520a0

Price, G. R. (1972). Extension of covariance selection mathematics. *Annals of Human Genetics*, *35*(4), 485–490. https://doi.org/10.1111/j.1469-1809.1957.tb01874.x

Pringle, J. (1951). On the Parallel Between Learning and Evolution. *Behaviour*, *3*(1), 174–214. https://doi.org/10.1163/156853951X00269

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/

Skinner, B. F. (1966). The phylogeny and ontogeny of behavior. Contingencies of reinforcement throw light on contingencies of survival in the evolution of behavior. *Science (New York, N.Y.)*, *153*(3741), 1205–1213. https://doi.org/10.1126/science.153.3741.1205

Skinner, B. F. (1981). Selection by consequences. *Science (New York, N.Y.)*, *213*(4507), 501–504.

Staddon, J. E. R. (2020). The Role of Theory in Behavior Analysis: A Response to Unfinished Business, Travis Thompson's Review of Staddon's New Behaviorism (2nd edition). *The Psychological Record*, 1–7. https://doi.org/10.1007/s40732-020-00409-y

Strand, P. S., Robinson, M. J. F., Fiedler, K. R., Learn, R., & Anselme, P. (2022). Quantifying the instrumental and noninstrumental underpinnings of Pavlovian responding with the Price equation. *Psychonomic Bulletin & Review*, *29*, 1295–1306. https://doi.org/10.3758/s13423-021-02047-z

Thorndike, E. L. (1900). The associative processes in animals. *Biological Lectures from the Marine Biological Laboratory of Woods Holl, 1899*, 69–91.

Timberlake, W., & Allison, J. (1974). Response deprivation: An empirical approach to instrumental performance. *Psychological Review*, *81*(2), 146–164. https://doi.org/10.1037/h0036101

Zentall, T. R. (2016). Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology. Animal Learning and Cognition*, *42*(1), 1–14. https://doi.org/10.1037/xan0000085

720